Title

stteffects ipwra — Survival-time inverse-probability-weighted regression adjustment

Description

`stteffects ipwra` estimates the average treatment effect (ATE), the average treatment effect on the treated (ATET), and the potential-outcome means (POMs) from observational survival-time data by inverse-probability-weighted regression adjustment (IPWRA). IPWRA estimators use missingness-adjusted regression coefficients to compute averages of treatment-level predicted outcomes. Contrasts of these averages estimate the treatment effects. `stteffects ipwra` offers several choices for the functional forms of the outcome model, of the treatment model, and of the optional time-to-censoring model. Binary and multivalued treatments are accommodated.

See [CAUSAL] **stteffects intro** for an overview of estimating treatment effects from observational survival-time data.

Quick start

Specify `time` as observed failure time and `fail` as failure indicator

```
stset time, failure(fail)
```

ATE of binary treatment `treat2` estimated by IPWRA using a Weibull model for `time` on `x1` and `x2` and a logistic model for `treat2` on `x1` and `w`

```
stteffects ipwra (x1 x2) (treat2 x1 w)
```

Same as above, but estimate the ATET

```
stteffects ipwra (x1 x2) (treat2 x1 w), atet
```

Gamma model for time and probit model for `treat2`

```
stteffects ipwra (x1 x2, gamma) (treat2 x1 w, probit)
```

ATE for each level of three-valued treatment `treat3`

```
stteffects ipwra (x1 x2) (treat3 x1 w)
```

Same as above, and specify that `treat3` = 3 is the control level using the value label "MyControl" for 3

```
stteffects ipwra (x1 x2) (treat3 x1 w), control("MyControl")
```

ATE of `treat2` estimated by IPWRA using a Weibull model for `time` on `x1` and `x2`, a logistic model for `treat2` on `x1` and `w`, and a Weibull model for the time to censoring with covariates `x1` and `x2`

```
stteffects ipwra (x1 x2) (treat2 x1 w) (x1 x2)
```

Gamma model for time, probit model for `treat2`, and gamma model for censoring

```
stteffects ipwra (x1 x2, gamma) (treat2 x1 w, probit) (x1 x2, gamma)
```

Menu

Statistics > Causal inference/treatment effects > Survival outcomes > Regression adjustment with IPW

Syntax

```
stteffects ipwra (omvarlist [ , omoptions ]) (tvar tmvarlist [ , tmoptions ])
    [ (cmvarlist [ , cmoptions ]) ] [ if ] [ in ] [ , stat options ]
```

omvarlist specifies the variables that predict the survival-time variable in the outcome model.

tvar must contain integer values representing the treatment levels.

tmvarlist specifies the variables that predict treatment assignment in the treatment model.

cmvarlist specifies the variables that predict censoring in the censoring model.

omoptions	Description
Model	
weibull	Weibull; the default
exponential	exponential
gamma	two-parameter gamma
lnormal	lognormal
ancillary(*avarlist* [, noconstant])	specify variables used to model ancillary parameter
noconstant	suppress constant from outcome model

tmoptions	Description
Model	
logit	logistic treatment model; the default
probit	probit treatment model
hetprobit(*varlist*)	heteroskedastic probit treatment model
noconstant	suppress constant from treatment model

cmoptions	Description
Model	
weibull	Weibull; the default
exponential	exponential
gamma	two-parameter gamma
lnormal	lognormal
ancillary(*avarlist* [, noconstant])	specify variables used to model ancillary parameter
noconstant	suppress constant from censoring model

stat	Description
Stat	
ate	estimate average treatment effect in population; the default
atet	estimate average treatment effect on the treated
pomeans	estimate potential-outcome means

options	Description
SE/Robust	
vce(*vcetype*)	*vcetype* may be robust, cluster *clustvar*, bootstrap, or jackknife
Reporting	
level(#)	set confidence level; default is level(95)
aequations	display auxiliary-equation results
noshow	do not show st setting information
display_options	control columns and column formats, row spacing, line width, display of omitted variables and base and empty cells, and factor-variable labeling
Maximization	
maximize_options	control the maximization process; seldom used
iterinit(#)	specify starting-value iterations; seldom used
Advanced	
pstolerance(#)	set tolerance for the overlap assumption
osample(*newvar*)	identify observations that violate the overlap assumption
control(# | *label*)	specify the level of *tvar* that is the control
tlevel(# | *label*)	specify the level of *tvar* that is the treatment
coeflegend	display legend instead of statistics

You must stset your data before using stteffects; see [ST] **stset**.

omvarlist, *tmvarlist*, *cmvarlist*, and *avarlist* may contain factor variables; see [U] **11.4.3 Factor variables**.

bootstrap, by, collect, jackknife, and statsby are allowed; see [U] **11.1.10 Prefix commands**.

Weights are not allowed with the bootstrap prefix; see [R] **bootstrap**.

fweights, iweights, and pweights may be specified using stset; see *Weights* under *Remarks and examples* in [ST] **stset**. However, weights may not be specified if you are using the bootstrap prefix.

coeflegend does not appear in the dialog box.

See [U] **20 Estimation and postestimation commands** for more capabilities of estimation commands.

Options

Model

ancillary(*avarlist* [, noconstant]) specifies the variables used to model the ancillary parameter. By default, the ancillary parameter does not depend on covariates. Specifying ancillary(*avarlist*, noconstant) causes the constant to be suppressed in the model for the ancillary parameter.

`ancillary()` may be specified for the model for survival-time outcome, for the model for the censoring variable, or for both. If `ancillary()` is specified for both, the varlist used for each model may be different.

`noconstant`; see [R] **Estimation options**.

Stat

stat is one of three statistics: `ate`, `atet`, or `pomeans`. `ate` is the default.

`ate` specifies that the average treatment effect be estimated.

`atet` specifies that the average treatment effect on the treated be estimated.

`pomeans` specifies that the potential-outcome means for each treatment level be estimated.

SE/Robust

`vce(`*vcetype*`)` specifies the type of standard error reported, which includes types that are robust to some kinds of misspecification (`robust`), that allow for intragroup correlation (`cluster` *clustvar*), and that use bootstrap or jackknife methods (`bootstrap`, `jackknife`); see [R] ***vce_option***.

Reporting

`level(#)`; see [R] **Estimation options**.

`aequations` specifies that the results for the outcome-model or treatment-model parameters be displayed. By default, the results for these auxiliary parameters are not displayed.

`noshow` prevents `stteffects ipwra` from showing the key st variables. This option is rarely used because most people type `stset, show` or `stset, noshow` to permanently set whether they want to see these variables mentioned at the top of the output of every st command; see [ST] **stset**.

display_options: `noci`, `nopvalues`, `noomitted`, `vsquish`, `noemptycells`, `baselevels`, `allbaselevels`, `nofvlabel`, `fvwrap(#)`, `fvwrapon(`*style*`)`, `cformat(%`*fmt*`)`, `pformat(%`*fmt*`)`, `sformat(%`*fmt*`)`, and `nolstretch`; see [R] **Estimation options**.

Maximization

maximize_options: `iterate(#)`, [`no`]`log`, and `from(`*init_specs*`)`; see [R] **Maximize**. These options are seldom used.

init_specs is one of

matname [`, skip copy`]

[`, #` ...]`, copy`

`iterinit(#)` specifies the maximum number of iterations used to calculate the starting values. This option is seldom used.

Advanced

`pstolerance(#)` specifies the tolerance used to check the overlap assumption. The default value is `pstolerance(1e-5)`. `stteffects` will exit with an error if an observation has an estimated propensity score smaller than that specified by `pstolerance()`.

`osample(`*newvar*`)` specifies that indicator variable *newvar* be created to identify observations that violate the overlap assumption.

`control(#|label)` specifies the level of *tvar* that is the control. The default is the first treatment level. You may specify the numeric level # (a nonnegative integer) or the label associated with the numeric level. `control()` may not be specified with the statistic `pomeans`. `control()` and `tlevel()` may not specify the same treatment level.

`tlevel(#|label)` specifies the level of *tvar* that is the treatment for the statistic `atet`. The default is the second treatment level. You may specify the numeric level # (a nonnegative integer) or the label associated with the numeric level. `tlevel()` may only be specified with statistic `atet`. `tlevel()` and `control()` may not specify the same treatment level.

The following option is available with `stteffects` but is not shown in the dialog box:

`coeflegend`; see [R] **Estimation options**.

Remarks and examples

stata.com

If you are not familiar with the framework for treatment-effects estimation from observational survival-time data, please see [CAUSAL] **stteffects intro**.

IPWRA estimators use estimated weights to obtain missingness-adjusted outcome-regression parameters. The missingness-adjusted outcome-regression parameters are used to compute averages of treatment-level predicted outcomes. Contrasts of these averages estimate the treatment effects.

The estimated weights account for the missing potential outcome and, optionally, for data lost to censoring. The weights are estimated using a treatment-assignment model and, optionally, a model for the censoring time. A term in the estimator for the outcome-regression parameters accounts for data lost to censoring when estimated weights are not used.

There are two versions of the IPWRA estimator because there are two methods of accounting for the data lost to censoring.

1. IPWRA estimators that adjust for censoring by including a term in the likelihood function for the outcome-model parameters are known as likelihood-adjusted-censoring IPWRA (LAC-IPWRA) estimators.
2. IPWRA estimators that adjust for censoring by weighting the likelihood function for the outcome-model parameters by estimated inverse-probability-of-censoring weights are known as weighted-adjusted-censoring IPWRA (WAC-IPWRA) estimators.

The LAC-IPWRA estimators require fewer assumptions than the WAC-IPWRA estimators. Outlining the steps performed by LAC-IPWRA and WAC-IPWRA estimators allows us to be more specific about the tradeoffs between the estimators.

LAC-IPWRA estimators use a three-step approach to estimating treatment effects:

1. Estimate the parameters of a treatment-assignment model and compute inverse-probability-of-treatment weights.
2. Obtain the treatment-specific predicted mean outcomes for each subject by using the weighted maximum likelihood estimators. Estimated inverse-probability-of-treatment weights are used to weight the maximum likelihood estimator. A term in the likelihood function adjusts for right-censored survival times.
3. Compute the means of the treatment-specific predicted mean outcomes. Contrasts of these averages provide the estimates of the ATEs. By restricting the computations of the means to the subset of treated subjects, we can obtain the ATETs.

WAC-IPWRA estimators use a four-step approach to estimating treatment effects:

1. Estimate the parameters of a treatment-assignment model and compute inverse-probability-of-treatment weights.
2. Estimate the parameters of a time-to-censoring model and compute inverse-probability-of-censoring weights.
3. Obtain the treatment-specific predicted mean outcomes for each subject by using the weighted maximum likelihood estimators. Estimated inverse-probability-of-treatment weights and inverse-probability-of-censoring weights are used to weight the maximum likelihood estimator. The inverse-probability-of-censoring weights account for right-censored survival times.
4. Compute the means of the treatment-specific predicted mean outcomes. Contrasts of these averages provide the estimates of the ATEs. By restricting the computations of the means to the subset of treated subjects, we can obtain the ATETs.

The WAC-IPWRA estimators require that the censoring time be random and that the time-to-censoring model be well specified. The implemented WAC-IPWRA estimators also require that the time-to-censoring process not vary by treatment level. The LAC-IPWRA estimators do not require these extra assumptions because they use a likelihood term instead of weights to adjust for the data lost to censoring.

Here we note only a few entry points to the vast literature on estimators that combine IPW and RA methods. Hirano, Imbens, and Ridder (2003), Imbens (2000, 2004), Imbens and Wooldridge (2009), Rosenbaum and Rubin (1983), Robins and Rotnitzky (1995, 2006), Robins, Rotnitzky, and Zhao (1995), Wooldridge (2002, 2007), Cameron and Trivedi (2005, chap. 25), Wooldridge (2010, chap. 21), and Vittinghoff et al. (2012, chap. 9) provide excellent general introductions to estimating ATEs and to the IPWRA estimators in particular.

Like `streg` and other survival-time commands, `stteffects ipwra` uses the outcome variable and the failure indicator computed by, and optionally weights specified with, `stset`. `stteffects ipwra` is not appropriate for data with time-varying covariates, also known as multiple-record survival-time data, or for delayed-entry data.

▷ Example 1: Estimating the ATE by LAC-IPWRA

Suppose we wish to study the effect of smoking on the time to a second heart attack among women aged 45–55 years. In our fictional `sheart` dataset, `atime` is the observed time in years to a second heart attack or censoring, and `fail` is the 0/1 indicator that a second heart attack was observed. (When `fail` is 1, `atime` records the time to the second heart attack; when `fail` is 0, `atime` records a censored observation of the time to a second heart attack.) We previously `stset` these data; see *A quick tour of the estimators* in [CAUSAL] **stteffects intro**.

The treatment, smoking, is stored in the 0/1 indicator `smoke`. These data also contain age at the time of the first heart attack (`age`), and indices of the level of exercise (`exercise`), diet quality (`diet`), and education (`education`) prior to the first heart attack.

We can use `stteffects ipwra` to estimate the ATE. We model the mean survival time using the default Weibull model, controlling for `age`, `exercise`, `diet`, and `education`. We model treatment assignment using the default logit model with covariates `age`, `exercise`, and `education`. We do not specify a time-to-censoring model so that we obtain the LAC estimator.

```
. use https://www.stata-press.com/data/r18/sheart
(Time to second heart attack (fictional))
. stteffects ipwra (age exercise diet education) (smoke age exercise education)
        Failure _d: fail
  Analysis time _t: atime
Iteration 0:  EE criterion =  7.439e-15
Iteration 1:  EE criterion =  1.756e-26
Survival treatment-effects estimation           Number of obs     =      2,000
Estimator      : IPW regression adjustment
Outcome model  : Weibull
Treatment model: logit
Censoring model: none
```

_t	Coefficient	Robust std. err.	z	P>|z|	[95% conf.	interval]
ATE						
smoke (Smoker vs Nonsmoker)	-1.591874	.4837332	-3.29	0.001	-2.539973	-.643774
POmean						
smoke Nonsmoker	4.214263	.2598689	16.22	0.000	3.704929	4.723597

When every woman smoked in the population of women aged 45–55 years who have had a heart attack, the average time to a second heart attack is estimated to be 1.59 years less than when no women in the population of interest smoked. The estimated average time to a second heart attack when no women in the population of interest smoked is 4.21 years.

The ratio of the ATE to the control-level potential-outcome mean (POM) measures the importance of the effect. In this example, when all women smoked, the time to the second heart attack falls by an estimated 38% relative to the case in which no women smoked. See example 3 in [CAUSAL] **stteffects ra** for an example that uses `nlcom` to compute a point estimate and a confidence interval for this ratio.

◁

Example 2: Different outcome and treatment models

Instead of a Weibull model for the outcome model, we could have used an exponential, a gamma, or a lognormal model. Instead of a logit model for the treatment assignment, we could have used a probit or a heteroskedastic probit model. This example uses a gamma model for the outcome and a probit model for the treatment assignment.

```
. stteffects ipwra (age exercise diet education, gamma)
> (smoke age exercise education, probit)
        Failure _d: fail
  Analysis time _t: atime
Iteration 0:  EE criterion =  2.644e-13
Iteration 1:  EE criterion =  2.153e-23
Survival treatment-effects estimation           Number of obs     =      2,000
Estimator      : IPW regression adjustment
Outcome model  : gamma
Treatment model: probit
Censoring model: none
```

_t	Coefficient	Robust std. err.	z	P>|z|	[95% conf.	interval]
ATE						
smoke (Smoker vs Nonsmoker)	-1.387303	.4786032	-2.90	0.004	-2.325348	-.4492583
POmean						
smoke Nonsmoker	3.97986	.2258474	17.62	0.000	3.537207	4.422512

The estimated ATE of -1.39 and control-level POM of 3.98 are similar to the values of -1.59 and 4.21 that we obtained in example 1.

Example 3: Estimating the ATE by WAC-IPWRA

Rather than using LAC, we may want to specify a time-to-censoring model. We now use `stteffects ipwra` to estimate the ATE by WAC-IPWRA. We use the same specification of the outcome and treatment models that we used in example 1. However, now we specify a time-to-censoring model, using the default Weibull model with covariates `age`, `exercise`, `diet`, and `education`.

```
. stteffects ipwra (age exercise diet education) (smoke age exercise education)
> (age exercise diet education)
        Failure _d: fail
  Analysis time _t: atime
Iteration 0:  EE criterion =  3.673e-18
Iteration 1:  EE criterion =  3.383e-31
Survival treatment-effects estimation           Number of obs     =      2,000
Estimator      : IPW regression adjustment
Outcome model  : Weibull
Treatment model: logit
Censoring model: Weibull
```

_t	Coefficient	Robust std. err.	z	P>|z|	[95% conf.	interval]
ATE						
smoke (Smoker vs Nonsmoker)	-2.285057	.7318456	-3.12	0.002	-3.719448	-.8506656
POmean						
smoke Nonsmoker	4.385841	.6427521	6.82	0.000	3.12607	5.645612

The estimated ATE of -2.29 differs from the ATE of -1.59 estimated by LAC-IPWRA, but the estimates of the control-level POM are similar between the two models: 4.39 for the WAC compared with 4.21 for the LAC.

◁

▷ Example 4: Estimating the ATET by LAC-IPWRA

Intuitively, the ATET measures the effect of the treatment on an at-risk subpopulation. Sometimes the subpopulation that gets the treatment defines such an at-risk subpopulation. The ATET has the added benefit that it can be estimated under weaker conditions than the ATE; see *Assumptions and tradeoffs* under *Remarks and examples* in [CAUSAL] **stteffects intro**.

```
. stteffects ipwra (age exercise diet education)
> (smoke age exercise education), atet
       Failure _d: fail
  Analysis time _t: atime
Iteration 0:  EE criterion =  2.815e-19
Iteration 1:  EE criterion =  6.331e-31
Survival treatment-effects estimation           Number of obs     =      2,000
Estimator      : IPW regression adjustment
Outcome model  : Weibull
Treatment model: logit
Censoring model: none
```

_t	Coefficient	Robust std. err.	z	P>|z|	[95% conf.	interval]
ATET						
smoke (Smoker vs Nonsmoker)	-1.775107	.3437506	-5.16	0.000	-2.448846	-1.101368
POmean						
smoke Nonsmoker	4.062424	.2779877	14.61	0.000	3.517578	4.60727

When all women in the subpopulation smoked, the average time to a second heart attack is estimated to be 1.78 years less than when no women in the subpopulation of interest smoked. If no women in the subpopulation of interest smoked, the average time to a second heart attack is 4.06 years.

◁

Stored results

`stteffects ipwra` stores the following in `e()`:

Scalars

`e(N)`	number of observations
`e(n`*j*`)`	number of observations for treatment level *j*
`e(N_clust)`	number of clusters
`e(k_eq)`	number of equations in `e(b)`
`e(k_levels)`	number of levels in treatment variable
`e(treated)`	level of treatment variable defined as treated
`e(control)`	level of treatment variable defined as control
`e(converged)`	`1` if converged, `0` otherwise

Macros

`e(cmd)`	`stteffects`
`e(cmdline)`	command as typed
`e(dead)`	`_d`
`e(depvar)`	`_t`
`e(tvar)`	name of treatment variable

e(subcmd)	ipwra
e(omodel)	outcome model: weibull, exponential, gamma, or lognormal
e(tmodel)	treatment model: logit, probit, or hetprobit
e(cmodel)	censoring model: weibull, exponential, gamma, or lognormal (if specified)
e(stat)	statistic estimated: ate, atet, or pomeans
e(wtype)	weight type
e(wexp)	weight expression
e(title)	title in estimation output
e(clustvar)	name of cluster variable
e(tlevels)	levels of treatment variable
e(vce)	*vcetype* specified in vce()
e(vcetype)	title used to label Std. err.
e(properties)	b V
e(estat_cmd)	program used to implement estat
e(predict)	program used to implement predict
e(marginsnotok)	predictions disallowed by margins
e(asbalanced)	factor variables fvset as asbalanced
e(asobserved)	factor variables fvset as asobserved

Matrices

e(b)	coefficient vector
e(V)	variance–covariance matrix of the estimators

Functions

e(sample)	marks estimation sample

In addition to the above, the following is stored in r():

Matrices

r(table)	matrix containing the coefficients with their standard errors, test statistics, p-values, and confidence intervals

Note that results stored in r() are updated when the command is replayed and will be replaced when any r-class command is run after the estimation command.

Methods and formulas

Methods and formulas are presented under the following headings:

Introduction

This section presents the methods and formulas used by the estimators implemented in stteffects ra, stteffects wra, stteffects ipw, and stteffects ipwra. This section assumes that you are familiar with the concepts and intuition from the estimators discussed in [CAUSAL] **teffects intro advanced**.

Each of the estimators implemented in stteffects has a multistep logic but is implemented as one step by simultaneously solving the estimating equations that define each step. This one-step estimating-equation approach provides consistent point estimates and a consistent variance–covariance of the estimator (VCE); see Newey (1984), Wooldridge (2010), and Drukker (2014).

Survival-time treatment-effects estimators handle two types of missing data. First, only one of the potential outcomes is observed, as is standard in causal inference. Second, the potential outcome for the received treatment may be censored. The data missing because of censoring may be handled by an outcome model, a censoring model, or both, just like the data missing due to observing only one potential outcome.

❑ Technical note

Delayed entry would be a third type of missing data. The left-truncation process caused by delayed entry would also need to be modeled to estimate ATE parameters. The estimators implement in `stteffects` do not allow for delayed entry because they do not have a method for modeling how the left-truncation process selects the sample, conditional on the covariates.

❑

All the implemented estimators are combinations of regression-adjustment (RA) and inverse-probability-weighted (IPW) techniques. RA estimators use an outcome model to account for the missing potential outcome and for censoring. IPW estimators use models for treatment assignment and censoring to construct weights that account for the missing potential outcome and for censoring.

The remainder of this section provides technical details about how the estimators in `stteffects` were implemented. We provide details only for the two-treatment-level case to simplify the formulas. We provide outlines for how the extensions to the multiple-treatment-level case were implemented.

Regression-adjusted estimators

We begin with the RA estimators implemented in `stteffects ra`. The RA estimators have the following logic:

RA1. For each treatment level $\tau \in \{0, 1\}$, estimate by maximum likelihood (ML) the parameters $\boldsymbol{\beta}_\tau$ of a parametric model for the survival-time outcome t in which $F(t|\mathbf{x}, \tau, \boldsymbol{\beta}_\tau)$ is the distribution of t conditional on covariates $\mathbf{x}$ and treatment level τ. Denote the estimates $\boldsymbol{\beta}_\tau$ by $\widehat{\boldsymbol{\beta}}_{\mathrm{ra},\tau}$.

RA2. Use the estimated $\widehat{\boldsymbol{\beta}}_{\mathrm{ra},\tau}$ and the functional form implied by $F(t|\mathbf{x}, \tau, \boldsymbol{\beta}_\tau)$ to estimate the mean survival time, conditional on $\mathbf{x}$ and treatment level τ, for each sample observation, denoted by $\widehat{E}(t_i|\mathbf{x}_i, \tau, \widehat{\boldsymbol{\beta}}_{\mathrm{ra},\tau})$. Conditional independence of the treatment and the survival-time potential outcomes ensures that $E(t|\mathbf{x}, \tau, \boldsymbol{\beta}_\tau) = E(t_\tau|\mathbf{x}, \boldsymbol{\beta}_\tau)$, where t_τ is the potential survival-time outcome corresponding to treatment level τ. Under correct model specification, sample averages of $\widehat{E}(t_i|\mathbf{x}_i, \tau, \widehat{\boldsymbol{\beta}}_{\mathrm{ra},\tau})$ consistently estimate the POM for treatment level τ, denoted by POM_τ.

RA3. A contrast of the estimated POMs estimates the ATE.

If estimating an ATET, step RA2 is modified to use only the treated observations when estimating the POMs. A contrast of these POMs then estimates the ATET.

The contribution of the ith observation to the log likelihood that is maximized in step RA1 is

$$L_{\mathrm{ra}}(t_i, \mathbf{x}_i, \tau, \widehat{\boldsymbol{\beta}}_{\mathrm{ra},\tau}) = \varpi_i(\tau_i == \tau) \Big[(1 - c_i) \ln\{f(t_i|\mathbf{x}_i, \tau, \widehat{\boldsymbol{\beta}}_{\mathrm{ra},\tau})\} + c_i \ln\{1 - F(t_i|\mathbf{x}_i, \tau, \widehat{\boldsymbol{\beta}}_{\mathrm{ra},\tau})\}\Big] \tag{1}$$

where ϖ_i is the observation-level weight, c_i is the 0/1 indicator for whether the survival-time observation on person i was censored, and $f(t_i|\mathbf{x}_i, \tau, \widehat{\boldsymbol{\beta}}_{\mathrm{ra},\tau})$ is the density corresponding to distribution $F(t_i|\mathbf{x}_i, \tau, \widehat{\boldsymbol{\beta}}_{\mathrm{ra},\tau})$. The first term inside the curly braces in (1) accounts for the noncensored observations, and the second term inside the curly braces accounts for the censored observations.

The RA estimators for the POMs simultaneously solve estimating equations (2a) through (2d) for $\widehat{\boldsymbol{\beta}}_{\mathrm{ra},0}$, $\widehat{\boldsymbol{\beta}}_{\mathrm{ra},1}$, $\widehat{\mathrm{POM}}_{\mathrm{ra},0}$, and $\widehat{\mathrm{POM}}_{\mathrm{ra},1}$.

$$1/N \sum_{i=1}^{N} \mathbf{s}_{\mathrm{ra}}(t_i, \mathbf{x}_i, 0, \widehat{\boldsymbol{\beta}}_{\mathrm{ra},0}, F) = \mathbf{0} \tag{2a}$$

$$1/N \sum_{i=1}^{N} \mathbf{s}_{\mathrm{ra}}(t_i, \mathbf{x}_i, 1, \widehat{\boldsymbol{\beta}}_{\mathrm{ra},1}, F) = \mathbf{0} \tag{2b}$$

$$1/N \sum_{i=1}^{N} \varpi_i \left\{ \widehat{E}(t_i|\mathbf{x}_i, \tau = 0, \widehat{\boldsymbol{\beta}}_{\mathrm{ra},0}) - \widehat{\mathrm{POM}}_{\mathrm{ra},0} \right\} = 0 \tag{2c}$$

$$1/N \sum_{i=1}^{N} \varpi_i \left\{ \widehat{E}(t_i|\mathbf{x}_i, \tau = 1, \widehat{\boldsymbol{\beta}}_{\mathrm{ra},1}) - \widehat{\mathrm{POM}}_{\mathrm{ra},1} \right\} = 0 \tag{2d}$$

where

$\mathbf{s}_{\mathrm{ra}}(t_i, \mathbf{x}_i, 0, \widehat{\boldsymbol{\beta}}_{\mathrm{ra},0}, F) = \frac{\partial L_{\mathrm{ra}}(t_i, \mathbf{x}_i, 0, \widehat{\boldsymbol{\beta}}_{\mathrm{ra},0})}{\partial \widehat{\boldsymbol{\beta}}_{\mathrm{ra},0}}$ is the vector of score equations from the ML estimator for $\widehat{\boldsymbol{\beta}}_{\mathrm{ra},0}$ based on survival-time model F,

$\mathbf{s}_{\mathrm{ra}}(t_i, \mathbf{x}_i, 1, \widehat{\boldsymbol{\beta}}_{\mathrm{ra},1}, F) = \frac{\partial L_{\mathrm{ra}}(t_i, \mathbf{x}_i, 1, \widehat{\boldsymbol{\beta}}_{\mathrm{ra},1})}{\partial \widehat{\boldsymbol{\beta}}_{\mathrm{ra},1}}$ is the vector of score equations from the ML estimator for $\widehat{\boldsymbol{\beta}}_{\mathrm{ra},1}$ based on survival-time model F,

$\widehat{E}(t_i|\mathbf{x}_i, \tau = 0, \widehat{\boldsymbol{\beta}}_{\mathrm{ra},0})$ is the predicted mean survival time assuming treatment level 0 for observation i conditional on $\mathbf{x}_i$, and

$\widehat{E}(t_i|\mathbf{x}_i, \tau = 1, \widehat{\boldsymbol{\beta}}_{\mathrm{ra},1})$ is the predicted mean survival time assuming treatment level 1 for observation i conditional on $\mathbf{x}_i$.

The ATE is estimated by replacing (2d) with

$$1/N \sum_{i=1}^{N} \varpi_i \left\{ \widehat{E}(t_i|\mathbf{x}_i, \tau = 1, \widehat{\boldsymbol{\beta}}_{\mathrm{ra},1}) - \widehat{\mathrm{POM}}_{\mathrm{ra},0} - \widehat{\mathrm{ATE}}_{\mathrm{ra}} \right\} = 0 \tag{3}$$

and the ATET is estimated by replacing (2c) and (3) with

$$1/N_1 \sum_{i=1}^{N} \varpi_i(\tau_i == 1) \left\{ \widehat{E}(t_i|\mathbf{x}_i, \tau = 0, \widehat{\boldsymbol{\beta}}_{\mathrm{ra},0}) - \widehat{\mathrm{POM}}_{\mathrm{ra,cot},0} \right\} = 0$$

$$1/N_1 \sum_{i=1}^{N} \varpi_i(\tau_i == 1) \left\{ \widehat{E}(t_i|\mathbf{x}_i, \tau = 1, \widehat{\boldsymbol{\beta}}_{\mathrm{ra},1}) - \widehat{\mathrm{POM}}_{\mathrm{ra,cot},0} - \widehat{\mathrm{ATET}}_{\mathrm{ra}} \right\} = 0$$

where $N_1 = \sum_{i=1}^{N}(t_i == 1)$ and $\widehat{\text{POM}}_{\text{ra,cot},0}$ is the estimated conditional-on-treatment POM for treatment level 0.

Asymptotic standard errors for estimating equation estimators, also known as exactly identified generalized method of moments estimators, are standard in the literature; see Newey (1984), Newey and McFadden (1994), Tsiatis (2006), and Wooldridge (2010). These standard errors always have a robust structure and have been generalized to cluster–robust standard errors (see Wooldridge [2010]).

The score equations and the functional form for the predicted mean survival time depend on the model for survival-time outcome F. We provide these details below, under *Functional-form details*.

Weighted-adjusted-censoring assumptions

All estimators that permit you to model the time to censoring are subject to three assumptions:

1. The censoring time must be random.
2. The censoring time must be from a known distribution.
3. The distribution of the censoring time cannot vary by treatment level.

We call these three requirements the WAC assumptions. If the WAC assumptions are violated, you can use either an RA estimator or the LAC version of the IPWRA estimator.

❑ Technical note

We now describe how the observed survival-time outcome t is generated from the random censoring time t_c, the received treatment τ, and the potential-outcome survival times t_0 and t_1 under the WAC assumptions. First, each potential outcome is either censored or not censored.

$$\widetilde{t}_0 = t_c(t_0 \geq t_c) + t_0\{1 - (t_0 \geq t_c)\}$$
$$\widetilde{t}_1 = t_c(t_1 \geq t_c) + t_1\{1 - (t_1 \geq t_c)\}$$

Under the WAC assumptions, t_c is a random variable from a known distribution, and t_c does not vary by treatment level.

Next, the received treatment $\tau \in \{0, 1\}$ determines which, possibly censored, potential outcome is observed.

$$t = (1 - \tau)\widetilde{t}_0 + \tau\widetilde{t}_1$$

The 0/1 indicator for whether the observed t was censored, denoted by c, is given by

$$c = (1 - \tau)(t_0 \geq t_c) + \tau(t_1 \geq t_c)$$

❑

Weighted regression-adjusted estimators

As is standard in the survival literature, the RA estimators account for censored survival times by adding a term to the log-likelihood function for censored observations [see (1)]. In contrast, weighted regression-adjustment (WRA) estimators use weights to account for censored observations and are subject to the WAC assumptions.

Wooldridge (2007) and Lin (2000) derived estimators for the regression parameters that maximize a weighted objective function of the uncensored observations. Each observation-level weight is the inverse of the probability of not being censored. Like the RA estimators, the WRA estimators use averages of the predicted mean survival times to estimate treatment-effect parameters.

The WRA estimators have the following logic.

WRA1. Estimate by ML the parameters $\boldsymbol{\gamma}$ of a parametric survival-time model for the time to censoring t_c, in which $F_c(t_c|\mathbf{w},\boldsymbol{\gamma})$ is the distribution of t_c conditional on covariates $\mathbf{w}$. Note that the censoring process does not vary by treatment level and that we only observe t_c when the observed potential outcome was censored. Denote the estimates of $\boldsymbol{\gamma}$ by $\widehat{\boldsymbol{\gamma}}$.

WRA2. For each treatment level $\tau \in \{0,1\}$, estimate by weighted maximum likelihood (WML) the $\boldsymbol{\beta}_\tau$ parameters of a parametric survival-time model, denoted by $F(t|\mathbf{x},\tau,\boldsymbol{\beta}_\tau)$, where t is the survival-time outcome and $\mathbf{x}$ are the covariates. The weights are the inverse of the estimated probabilities of not being censored, $1/\{1-F_c(t_c|\mathbf{w},\widehat{\boldsymbol{\gamma}})\}$, and only the uncensored observations are used. Denote the estimates of $\boldsymbol{\beta}_\tau$ by $\widehat{\boldsymbol{\beta}}_{\mathrm{wra},\tau}$.

WRA3. Use the estimated $\widehat{\boldsymbol{\beta}}_{\mathrm{wra},\tau}$ and the functional form implied by $F(t|\mathbf{x},\tau,\boldsymbol{\beta}_\tau)$ to estimate the mean survival time, conditional on $\mathbf{x}$ and treatment level τ, for each sample observation, denoted by $\widehat{E}(t_i|\mathbf{x}_i,\tau,\widehat{\boldsymbol{\beta}}_{\mathrm{wra},\tau})$. Conditional independence of the treatment and the survival-time potential outcomes ensures that $E(t|\mathbf{x},\tau,\boldsymbol{\beta}_\tau) = E(t_\tau|\mathbf{x},\boldsymbol{\beta}_\tau)$, where t_τ is the potential survival-time outcome corresponding to treatment level τ. Under correct model specification, sample averages of $\widehat{E}(t_i|\mathbf{x}_i,\tau,\widehat{\boldsymbol{\beta}}_{\mathrm{wra},\tau})$ consistently estimate the POM for treatment level τ, denoted by POM_τ.

WRA4. A contrast of the estimated POMs estimates the ATE.

If estimating an ATET, step WRA3 is modified to use only the treated observations when estimating the POMs. A contrast of these POMs then estimates the ATET.

The contribution of the ith observation to the log likelihood that is maximized in step WRA1 is

$$L_{c,\mathrm{wra}}(t_i,\mathbf{w}_i,\widehat{\boldsymbol{\gamma}}) = \varpi_i\left[c_i \ln\{f_c(t_i|\mathbf{w}_i,\widehat{\boldsymbol{\gamma}})\} + (1-c_i)\ln\{1-F_c(t_i|\mathbf{w}_i,\widehat{\boldsymbol{\gamma}})\}\right] \tag{4}$$

where ϖ_i is the observation-level weight, c_i is the 0/1 indicator for whether the survival-time observation on person i was censored, t_i is the observed failure time, and $f_c(t_i|\mathbf{w}_i,\widehat{\boldsymbol{\gamma}})$ is the density corresponding to conditional time-to-censoring distribution $F_c(t_i|\mathbf{w}_i,\widehat{\boldsymbol{\gamma}})$. When $c_i = 1$, t_i is the time to censoring. When $c_i = 0$, the censoring time is not observed; we only know that it is greater than the observed t_i. The first term accounts for the observations in which t_i is observed to be the censoring time, and the second term accounts for the observations in which the censoring time is greater than the observed t_i.

The contribution of the ith observation to the log likelihood that is maximized in step WRA2 is

$$L_{\mathrm{wra}}(t_i,\mathbf{x}_i,\tau,\widehat{\boldsymbol{\beta}}_{\mathrm{wra},\tau}) = \varpi_i(\tau_i == \tau)\left[\frac{(1-c_i)}{\{1-F_c(t_i|\mathbf{w}_i,\widehat{\boldsymbol{\gamma}})\}}\right]\ln\{f(t_i|\mathbf{x}_i,\tau,\widehat{\boldsymbol{\beta}}_{\mathrm{wra},\tau})\} \tag{5}$$

where $f(t_i|\mathbf{x}_i,\tau,\widehat{\boldsymbol{\beta}}_{\mathrm{wra},\tau})$ is the density corresponding to distribution $F(t_i|\mathbf{x}_i,\tau,\widehat{\boldsymbol{\beta}}_{\mathrm{wra},\tau})$. Equation (5) does not contain a term that adjusts for censoring; see (1) for a comparison. Rather, it uses inverse-probability weights to account for both the censored and the uncensored observations.

The WRA estimators for the POMs simultaneously solve estimating equations (6a) through (6e) for $\widehat{\boldsymbol{\gamma}}$, $\widehat{\boldsymbol{\beta}}_{\mathrm{wra},0}$, $\widehat{\boldsymbol{\beta}}_{\mathrm{wra},1}$, $\widehat{\mathrm{POM}}_{\mathrm{wra},0}$, and $\widehat{\mathrm{POM}}_{\mathrm{wra},1}$.

$$1/N\sum_{i=1}^{N}\mathbf{s}_{\text{wra}}(t_i,\mathbf{w}_i,\widehat{\boldsymbol{\gamma}},F_c)=\mathbf{0} \tag{6a}$$

$$1/N\sum_{i=1}^{N}\mathbf{s}_{\text{wra}}(t_i,\mathbf{x}_i,0,\widehat{\boldsymbol{\beta}}_{\text{wra},0},F)=\mathbf{0} \tag{6b}$$

$$1/N\sum_{i=1}^{N}\mathbf{s}_{\text{wra}}(t_i,\mathbf{x}_i,1,\widehat{\boldsymbol{\beta}}_{\text{wra},1},F)=\mathbf{0} \tag{6c}$$

$$1/N\sum_{i=1}^{N}\varpi_i\left\{\widehat{E}(t_i|\mathbf{x}_i,\tau=0,\widehat{\boldsymbol{\beta}}_{\text{wra},0})-\widehat{\text{POM}}_{\text{wra},0}\right\}=0 \tag{6d}$$

$$1/N\sum_{i=1}^{N}\varpi_i\left\{\widehat{E}(t_i|\mathbf{x}_i,\tau=1,\widehat{\boldsymbol{\beta}}_{\text{wra},1})-\widehat{\text{POM}}_{\text{wra},1}\right\}=0 \tag{6e}$$

where

$\mathbf{s}_{\text{wra}}(t_i,\mathbf{w}_i,\widehat{\boldsymbol{\gamma}},F_c)=\frac{\partial L_{c,\text{wra}}(t_i,\mathbf{w}_i,\widehat{\boldsymbol{\gamma}})}{\partial\widehat{\boldsymbol{\gamma}}}$ is the vector of score equations from the ML estimator for $\widehat{\boldsymbol{\gamma}}$ based on survival-time model F_c,

$\mathbf{s}_{\text{wra}}(t_i,\mathbf{x}_i,0,\widehat{\boldsymbol{\beta}}_{\text{wra},0},F)=\frac{\partial L(t_i,\mathbf{x}_i,0,\widehat{\boldsymbol{\beta}}_{\text{wra},0})}{\partial\widehat{\boldsymbol{\beta}}_{\text{wra},0}}$ is the vector of score equations from the WML estimator for $\widehat{\boldsymbol{\beta}}_{\text{wra},0}$ based on survival-time model F,

$\mathbf{s}_{\text{wra}}(t_i,\mathbf{x}_i,1,\widehat{\boldsymbol{\beta}}_{\text{wra},1},F)=\frac{\partial L(t_i,\mathbf{x}_i,1,\widehat{\boldsymbol{\beta}}_{\text{wra},1})}{\partial\widehat{\boldsymbol{\beta}}_{\text{wra},1}}$ is the vector of score equations from the WML estimator for $\widehat{\boldsymbol{\beta}}_{\text{wra},1}$ based on survival-time model F,

$\widehat{E}(t_i|\mathbf{x}_i,\tau=0,\widehat{\boldsymbol{\beta}}_{\text{wra},0})$ is the predicted mean survival time assuming treatment level 0 for observation i conditional on $\mathbf{x}_i$, and

$\widehat{E}(t_i|\mathbf{x}_i,\tau=1,\widehat{\boldsymbol{\beta}}_{\text{wra},1})$ is the predicted mean survival time assuming treatment level 1 for observation i conditional on $\mathbf{x}_i$.

The observation-level scores $\mathbf{s}_{\text{wra}}(t_i,\mathbf{x}_i,0,\widehat{\boldsymbol{\beta}}_{\text{wra},0},F)$ and $\mathbf{s}_{\text{wra}}(t_i,\mathbf{x}_i,1,\widehat{\boldsymbol{\beta}}_{\text{wra},1},F)$ also depend on c_i, $\mathbf{w}_i$, $\widehat{\boldsymbol{\gamma}}$, and F_c, but we ignored this dependence to simplify the notation.

The ATE is estimated by replacing (6e) with

$$1/N\sum_{i=1}^{N}\varpi_i\left\{\widehat{E}(t_i|\mathbf{x}_i,\tau=1,\widehat{\boldsymbol{\beta}}_{\text{wra},1})-\widehat{\text{POM}}_{\text{wra},0}-\widehat{\text{ATE}}_{\text{wra}}\right\}=0 \tag{7}$$

and the ATET is estimated by replacing (6e) and (7) with

$$1/N_1\sum_{i=1}^{N}\varpi_i(\tau_i==1)\left\{\widehat{E}(t_i|\mathbf{x}_i,\tau=0,\widehat{\boldsymbol{\beta}}_{\text{wra},0})-\widehat{\text{POM}}_{\text{wra,cot},0}\right\}=0$$

$$1/N_1\sum_{i=1}^{N}\varpi_i(\tau_i==1)\left\{\widehat{E}(t_i|\mathbf{x}_i,\tau=1,\widehat{\boldsymbol{\beta}}_{\text{wra},1})-\widehat{\text{POM}}_{\text{wra,cot},0}-\widehat{\text{ATET}}_{\text{wra}}\right\}=0$$

where $\widehat{\text{POM}}_{\text{wra,cot},0}$ is the estimated conditional-on-treatment POM.

Inverse-probability-weighted estimators

IPW estimators are weighted averages of the observed outcome. The weights correct for missing data due to unobserved potential outcomes and censoring. Each weight is the inverse of the probability that a given value is observed. Observed values that were not likely to be observed have higher weights.

When the outcome variable is never censored, the missing data are the unobserved potential outcome and an observation's weight is the inverse of a treatment probability. When the outcome may be censored, the censoring is an additional source of missing data. In this case, an observation's weight is the inverse of the joint probability that an observation is uncensored and has a particular treatment level.

To define this joint probability, the censoring time must be random. In practice, we make the WAC assumptions.

As is standard in the survival-time literature, we assume that the censoring-time process is independent of treatment assignment after conditioning on the covariates. This conditional independence assumption implies that the probability that observation i receives treatment level 1 and is not censored is the product of the probability that i gets treatment level 1 and the probability that i is not censored at time t_i, which we denote by

$$p(\mathbf{z}_i, \boldsymbol{\alpha})\{1 - F_c(t_i|\mathbf{w}_i, \boldsymbol{\gamma})\}$$

where

- $p(\mathbf{z}_i, \boldsymbol{\alpha})$ is the modeled probability that i gets treatment level 1, conditional on covariates $\mathbf{z}_i$ with parameters $\boldsymbol{\alpha}$, and
- $F_c(t_i|\mathbf{w}_i, \boldsymbol{\gamma})$ is the survival-time model for the censoring time, conditional on covariates $\mathbf{w}_i$ with parameters $\boldsymbol{\gamma}$, and evaluated at time t_i.

Bai, Tsiatis, and O'Brien (2013) formally derive these weights to control jointly for the missing potential outcome and censoring.

The IPW estimators have the following logic.

IPW1. Estimate by ML the parameters $\boldsymbol{\gamma}$ of a parametric survival-time model for the time to censoring, in which $F_c(t_c|\mathbf{w}, \boldsymbol{\gamma})$ is the distribution of censoring time, conditional on covariates $\mathbf{w}$. Denote the estimates of $\boldsymbol{\gamma}$ by $\widehat{\boldsymbol{\gamma}}$.

IPW2. Estimate by ML the parameters $\boldsymbol{\alpha}$ of a parametric model for the probability of treatment model $p(\mathbf{z}_i, \boldsymbol{\alpha})$. Denote the estimates of $\boldsymbol{\alpha}$ by $\widehat{\boldsymbol{\alpha}}$.

IPW3. Use the $\widehat{\boldsymbol{\gamma}}$ estimated in IPW1 and the $\widehat{\boldsymbol{\alpha}}$ estimated in IPW2 to construct inverse-probability weights by (8a) for treatment level 1 and by (8b) for treatment level 0.

$$\omega_{i,1} = \frac{(\tau_i == 1)(c_i == 0)}{[p(\mathbf{z}_i, \widehat{\boldsymbol{\alpha}})\{1 - F_c(t_i|\mathbf{w}_i, \widehat{\boldsymbol{\gamma}})\}]} \tag{8a}$$

$$\omega_{i,0} = \frac{(\tau_i == 0)(c_i == 0)}{[\{1 - p(\mathbf{z}_i, \widehat{\boldsymbol{\alpha}})\}\{1 - F_c(t_i|\mathbf{w}_i, \widehat{\boldsymbol{\gamma}})\}]} \tag{8b}$$

IPW4. Use the estimated weights to estimate each POM_τ by a weighted average of the uncensored observations on the observed potential outcome.

The contribution of the ith observation to the log likelihood that is maximized in step IPW1 is

$$L_{c,\text{ipw}}(t_i, \mathbf{w}_i, \widehat{\boldsymbol{\gamma}}) = \varpi_i \left[c_i \ln\{f_c(t_i|\mathbf{w}_i, \widehat{\boldsymbol{\gamma}})\} + (1 - c_i) \ln\{1 - F_c(t_i|\mathbf{w}_i, \widehat{\boldsymbol{\gamma}})\} \right]$$

where the definitions and intuition are as described after (4).

The contribution of the ith observation to the log likelihood that is maximized in step IPW2 is

$$L_{p,\text{ipw}}(\tau_i, \mathbf{z}_i, \widehat{\boldsymbol{\alpha}}) = \varpi_i \left[(\tau_i == 1) \ln\{p(\mathbf{z}_i, \widehat{\boldsymbol{\alpha}})\} + \{1 - (\tau_i == 1)\} \ln\{1 - p(\mathbf{z}_i, \widehat{\boldsymbol{\alpha}})\} \right]$$

where $p(\mathbf{z}_i, \widehat{\boldsymbol{\alpha}})$ is the model for the probability that i gets treatment level 1.

The IPW estimators for the POMs simultaneously solve estimating equations (9a) through (9d) for $\widehat{\boldsymbol{\gamma}}$, $\widehat{\boldsymbol{\alpha}}$, $\widehat{\text{POM}}_{\text{ipw},0}$, and $\widehat{\text{POM}}_{\text{ipw},1}$.

$$1/N \sum_{i=1}^{N} \mathbf{s}_{\text{ipw}}(t_i, \mathbf{w}_i, \widehat{\boldsymbol{\gamma}}, F_c) = \mathbf{0} \tag{9a}$$

$$1/N \sum_{i=1}^{N} \mathbf{s}_{\text{ipw}}(\tau_i, \mathbf{z}_i, \widehat{\boldsymbol{\alpha}}, p) = \mathbf{0} \tag{9b}$$

$$1/N \sum_{i=1}^{N} \varpi_i \omega_{i,0} \left(t_i - \widehat{\text{POM}}_{\text{ipw},0} \right) = 0 \tag{9c}$$

$$1/N \sum_{i=1}^{N} \varpi_i \omega_{i,1} \left(t_i - \widehat{\text{POM}}_{\text{ipw},1} \right) = 0 \tag{9d}$$

where

$\mathbf{s}_{\text{ipw}}(t_i, \mathbf{w}_i, \widehat{\boldsymbol{\gamma}}, F_c) = \frac{\partial L_{c,\text{ipw}}(t_i, \mathbf{w}_i, \widehat{\boldsymbol{\gamma}})}{\partial \widehat{\boldsymbol{\gamma}}}$ is the vector of score equations from the ML estimator for $\widehat{\boldsymbol{\gamma}}$ based on survival-time model F_c, and

$\mathbf{s}_{\text{ipw}}(\tau_i, \mathbf{z}_i, \widehat{\boldsymbol{\alpha}}, p) = \frac{\partial L_{p,\text{ipw}}(\tau_i, \mathbf{z}_i, \widehat{\boldsymbol{\alpha}})}{\partial \widehat{\boldsymbol{\alpha}}}$ is the vector of score equations from the ML estimator for $\widehat{\boldsymbol{\alpha}}$ based on probability model p.

The literature on IPW estimators discusses using normalized versus unnormalized weights, with normalized weights doing better in simulation studies; see Busso, DiNardo, and McCrary (2014) for example. The way that weights enter moment equations (9c) and (9d) implies that they are normalized, because the scale of the weights does not affect the estimates.

The estimated ATE is computed as

$$\widehat{\text{POM}}_{\text{ipw},1} - \widehat{\text{POM}}_{\text{ipw},0} = \widehat{\text{ATE}}_{\text{ipw}}$$

The estimated ATET uses weights

$$\omega_{i,\text{cot},1} = \frac{(\tau_i == 1)(c_i == 0)}{[\{1 - F_c(t_i|\mathbf{w}_i, \widehat{\boldsymbol{\gamma}})\}]} \tag{10a}$$

for treatment level 1 and

$$\omega_{i,\text{cot},0} = \frac{p(\mathbf{z}_i, \widehat{\boldsymbol{\alpha}})(\tau_i == 0)(c_i == 0)}{[\{1 - p(\mathbf{z}_i, \widehat{\boldsymbol{\alpha}})\}\{1 - F_c(t_i|\mathbf{w}_i, \widehat{\boldsymbol{\gamma}})\}]} \tag{10b}$$

for treatment level 0, and replaces (9c) and (9d) with

$$1/N_1 \sum_{i=1}^{N} \varpi_i \omega_{i,\mathrm{cot},0} \left(t_i - \widehat{\mathrm{POM}}_{\mathrm{ipw},\mathrm{cot},0}\right) = 0 \tag{11a}$$

$$1/N_1 \sum_{i=1}^{N} \varpi_i \omega_{i,\mathrm{cot},1} \left(t_i - \widehat{\mathrm{POM}}_{\mathrm{ipw},\mathrm{cot},1}\right) = 0 \tag{11b}$$

and then computes

$$\widehat{\mathrm{POM}}_{\mathrm{ipw},\mathrm{cot},1} - \widehat{\mathrm{POM}}_{\mathrm{ipw},\mathrm{cot},0} = \widehat{\mathrm{ATET}}_{\mathrm{ipw}}$$

These IPW estimators can be viewed as weighted IPW estimators and are thus related to those in Hirano, Imbens, and Ridder (2003).

Uncensored data

As mentioned, when the outcome variable is never censored, the missing data are the unobserved potential outcome and an observation's weight is the inverse of a treatment probability. In the never-censored case, the IPW estimators are identical to those implemented in `teffects ipw`; see *IPW estimators* under *Methods and formulas* in [CAUSAL] **teffects aipw**.

`stteffects ipw` computes the estimator for never-censored data when a censoring model is not specified and there are no censored observations in the sample. In the never-censored case, the following changes are made to the IPW estimator for the POMs and the ATE.

1. Step IPW1 is not performed.
2. The weights in (8a) and (8b) for the POMs and the ATE are replaced with (12a) for treatment level 1 and (12b) for treatment level 0.

$$\omega_{i,1} = \frac{(\tau_i == 1)}{p(\mathbf{z}_i, \widehat{\boldsymbol{\alpha}})} \tag{12a}$$

$$\omega_{i,0} = \frac{(\tau_i == 0)}{\{1 - p(\mathbf{z}_i, \widehat{\boldsymbol{\alpha}})\}} \tag{12b}$$

3. Only moment conditions (9b), (9c), and (9d) are used.

The following changes also are made to the IPW estimator for the ATET.

1. Step IPW1 is not performed.
2. The weights in (10a) and (10b) are replaced with (13a) for treatment level 1 and (13b) for treatment level 0.

$$\omega_{i,\mathrm{cot},1} = (\tau_i == 1) \tag{13a}$$

$$\omega_{i,\mathrm{cot},0} = \frac{p(\mathbf{z}_i, \widehat{\boldsymbol{\alpha}})(\tau_i == 0)}{\{1 - p(\mathbf{z}_i, \widehat{\boldsymbol{\alpha}})\}} \tag{13b}$$

3. Only moment conditions (9b), (11a), and (11b) are used.

Inverse-probability-weighted regression-adjustment estimators

IPWRA estimators are averages of treatment-specific predicted conditional means that were made using missingness-adjusted regression parameters. These estimators are Wooldridge's IPWRA for survival-time outcomes; see Wooldridge (2010, chap. 21) and Wooldridge (2007).

The censored observations can be handled either by weighting under the WAC assumptions to obtain the WAC-IPWRA estimator or by adding a term to the log-likelihood function (which we call likelihood-adjusted censoring) to obtain the LAC-IPWRA estimator. Correspondingly, there are two versions of formulas for the IPWRA estimator.

1. When a censoring model is specified, `stteffects ipwra` uses the formulas for the WAC-IPWRA estimator given in *Weighted-adjusted-censoring IPWRA*.
2. When a censoring model is not specified, `stteffects ipwra` uses the formulas for the LAC-IPWRA given in *Likelihood-adjusted-censoring IPWRA*, below.

The WAC-IPWRA estimator requires that some observations be censored and that the WAC assumptions hold; see *Weighted-adjusted-censoring assumptions*, above. The LAC-IPWRA estimator handles the case in which no observations are censored and requires the weaker independent censoring assumptions, which allows for fixed censoring times.

Weighted-adjusted-censoring IPWRA

When a censoring model is specified, `stteffects ipwra` uses the formulas for the WAC-IPWRA estimator to obtain the model-based weights that account for censoring. For notational conciseness and to reinforce its dependence on random censoring, we denote the WAC-IPWRA estimator by IPWRAR in lists and formulas. The WAC-IPWRA estimators have the following logic.

IPWRAR1. Estimate by ML the parameters $\boldsymbol{\gamma}$ of a parametric survival-time model for the time to censoring, in which $F_c(t_c|\mathbf{w},\boldsymbol{\gamma})$ is the censoring-time distribution, conditional on covariates $\mathbf{w}$. We denote the estimates of $\boldsymbol{\gamma}$ by $\widehat{\boldsymbol{\gamma}}$.

IPWRAR2. Estimate by ML the parameters $\boldsymbol{\alpha}$ of a parametric model for the probability of treatment model $p(\mathbf{z}_i,\boldsymbol{\alpha})$. We denote the estimates of $\boldsymbol{\alpha}$ by $\widehat{\boldsymbol{\alpha}}$.

IPWRAR3. For each treatment level $\tau \in \{0,1\}$, estimate by WML the parameters $\boldsymbol{\beta}_\tau$ of a parametric model for the survival-time outcome t, in which $F(t|\mathbf{x},\tau,\boldsymbol{\beta}_\tau)$ is the distribution of t conditional on covariates $\mathbf{x}$ and treatment level τ. For the ATE, the weights are those in equations (8a) and (8b). For the ATET, the weights are those in equations (10a) and (10b). We denote the estimates of $\boldsymbol{\beta}_{\text{ipwrar},\tau}$ by $\widehat{\boldsymbol{\beta}}_\tau$.

IPWRAR4. Use the estimated $\widehat{\boldsymbol{\beta}}_{\text{ipwrar},\tau}$ and the functional form implied by $F(t|\mathbf{x},\tau,\boldsymbol{\beta}_\tau)$ to estimate the mean survival time, conditional on $\mathbf{x}$ and treatment level τ, for each sample observation, denoted by $\widehat{E}(t_i|\mathbf{x}_i,\tau,\widehat{\boldsymbol{\beta}}_{\text{ipwrar},\tau})$. Conditional independence of the treatment and the survival-time potential outcomes ensures that $E(t|\mathbf{x},\tau,\boldsymbol{\beta}_\tau) = E(t_\tau|\mathbf{x},\boldsymbol{\beta}_\tau)$, where t_τ is the potential survival-time outcome corresponding to treatment level τ. Under correct model specification, sample averages of $\widehat{E}(t_i|\mathbf{x}_i,\tau,\widehat{\boldsymbol{\beta}}_{\text{ipwrar},\tau})$ consistently estimate the POM for treatment level τ, denoted by POM_τ.

The contribution of the ith observation to the log likelihood that is maximized in step IPWRAR1 is

$$L_{c,\text{ipwrar}}(t_i,\mathbf{w}_i,\widehat{\boldsymbol{\gamma}}) = \varpi_i\left[c_i \ln\{f_c(t_i|\mathbf{w}_i,\widehat{\boldsymbol{\gamma}})\} + (1-c_i)\ln\{1-F_c(t_i|\mathbf{w}_i,\widehat{\boldsymbol{\gamma}})\}\right]$$

where the definitions and intuition are as described after (4).

The contribution of the ith observation to the log likelihood that is maximized in step IPWRAR2 is

$$L_{p,\text{ipwrar}}(\tau_i, \mathbf{z}_i, \widehat{\boldsymbol{\alpha}}) = \varpi_i \left[(\tau_i == 1) \ln\{p(\mathbf{z}_i, \widehat{\boldsymbol{\alpha}})\} + \{1 - (\tau_i == 1)\} \ln\{1 - p(\mathbf{z}_i, \widehat{\boldsymbol{\alpha}})\}\right]$$

where $p(\mathbf{z}_i, \widehat{\boldsymbol{\alpha}})$ is the model for the probability that i gets treatment level 1.

The weights and the parameters in step IPWRAR3 used to estimate the ATE differ from those used to estimate the ATET. For the ATE, the contribution of the ith observation to the log likelihood that is maximized in step IPWRAR3 is

$$L_{\text{ipwrar}}(t_i, \mathbf{x}_i, \tau, \widehat{\boldsymbol{\beta}}_{\text{ipwrar,ate},\tau}) = \varpi_i \omega_{i,\tau} \ln\{f(t_i|\mathbf{x}_i, \tau, \widehat{\boldsymbol{\beta}}_{\text{ipwrar,ate},\tau})\}$$

where $\omega_{i,1}$ is given in (8a), $\omega_{i,0}$ is given in (8b), and $f(t_i|\mathbf{x}_i, \tau, \widehat{\boldsymbol{\beta}}_{\text{ipwrar,ate},\tau})$ is the density corresponding to distribution $F(t_i|\mathbf{x}_i, \tau, \widehat{\boldsymbol{\beta}}_{\text{ipwrar,ate},\tau})$. Like WRA, only the uncensored observations are used because the weights account for censoring.

The IPWRAR estimators for the POMs simultaneously solve estimating equations (14a) through (14f) for $\widehat{\boldsymbol{\gamma}}$, $\widehat{\boldsymbol{\alpha}}$, $\widehat{\boldsymbol{\beta}}_{\text{ipwrar,ate},0}$, $\widehat{\boldsymbol{\beta}}_{\text{ipwrar,ate},0}$, $\widehat{\text{POM}}_{\text{ipwrar},0}$, and $\widehat{\text{POM}}_{\text{ipwrar},1}$.

$$1/N \sum_{i=1}^{N} \mathbf{s}_{\text{ipwrar}}(t_i, \mathbf{w}_i, \widehat{\boldsymbol{\gamma}}, F_c) = \mathbf{0} \tag{14a}$$

$$1/N \sum_{i=1}^{N} \mathbf{s}_{\text{ipwrar}}(\tau_i, \mathbf{z}_i, \widehat{\boldsymbol{\alpha}}, p) = \mathbf{0} \tag{14b}$$

$$1/N \sum_{i=1}^{N} \mathbf{s}_{\text{ipwrar}}(t_i, \mathbf{x}_i, 0, \widehat{\boldsymbol{\beta}}_{\text{ipwrar,ate},0}, F) = \mathbf{0} \tag{14c}$$

$$1/N \sum_{i=1}^{N} \mathbf{s}_{\text{ipwrar}}(t_i, \mathbf{x}_i, 1, \widehat{\boldsymbol{\beta}}_{\text{ipwrar,ate},1}, F) = \mathbf{0} \tag{14d}$$

$$1/N \sum_{i=1}^{N} \varpi_i \left\{ \widehat{E}(t_i|\mathbf{x}_i, \tau = 0, \widehat{\boldsymbol{\beta}}_{\text{ipwrar,ate},0}) - \widehat{\text{POM}}_{\text{ipwrar},0} \right\} = 0 \tag{14e}$$

$$1/N \sum_{i=1}^{N} \varpi_i \left\{ \widehat{E}(t_i|\mathbf{x}_i, \tau = 1, \widehat{\boldsymbol{\beta}}_{\text{ipwrar,ate},1}) - \widehat{\text{POM}}_{\text{ipwrar},1} \right\} = 0 \tag{14f}$$

where

$\mathbf{s}_{\text{ipwrar}}(t_i, \mathbf{w}_i, \widehat{\boldsymbol{\gamma}}, F_c) = \frac{\partial L_{c,\text{ipwrar}}(t_i, \mathbf{w}_i, \widehat{\boldsymbol{\gamma}})}{\partial \widehat{\boldsymbol{\gamma}}}$ is the vector of score equations from the ML estimator for $\widehat{\boldsymbol{\gamma}}$ based on survival-time model F_c,

$\mathbf{s}_{\text{ipwrar}}(\tau_i, \mathbf{z}_i, \widehat{\boldsymbol{\alpha}}, p) = \frac{\partial L_{p,\text{ipwrar}}(\tau_i, \mathbf{z}_i, \widehat{\boldsymbol{\alpha}})}{\partial \widehat{\boldsymbol{\alpha}}}$ is the vector of score equations from the ML estimator for $\widehat{\boldsymbol{\alpha}}$ based on probability model p,

$\mathbf{s}_{\text{ipwrar}}(t_i, \mathbf{x}_i, 0, \widehat{\boldsymbol{\beta}}_{\text{ipwrar,ate},0}, F) = \frac{\partial L_{\text{ipwrar}}(t_i, \mathbf{x}_i, 0, \widehat{\boldsymbol{\beta}}_{\text{ipwrar,ate},0})}{\partial \widehat{\boldsymbol{\beta}}_{\text{ipwrar,ate},0}}$ is the vector of score equations from the ML estimator for $\widehat{\boldsymbol{\beta}}_{\text{ipwrar,ate},0}$ based on survival-time model F,

$\mathbf{s}_{\text{ipwrar}}(t_i, \mathbf{x}_i, 1, \widehat{\boldsymbol{\beta}}_{\text{ipwrar,ate},1}, F) = \frac{\partial L_{\text{ipwrar}}(t_i, \mathbf{x}_i, 1, \widehat{\boldsymbol{\beta}}_{\text{ipwrar,ate},1})}{\partial \widehat{\boldsymbol{\beta}}_{\text{ipwrar,ate},1}}$ is the vector of score equations from the ML estimator for $\widehat{\boldsymbol{\beta}}_{\text{ipwrar,ate},1}$ based on survival-time model F,

$\widehat{E}(t_i|\mathbf{x}_i, \tau = 0, \widehat{\boldsymbol{\beta}}_{\text{ipwrar,ate},0})$ is the predicted mean survival time assuming treatment level 0 for observation i conditional on $\mathbf{x}_i$, and

$\widehat{E}(t_i|\mathbf{x}_i, \tau = 1, \widehat{\boldsymbol{\beta}}_{\text{ipwrar,ate},1})$ is the predicted mean survival time assuming treatment level 1 for observation i conditional on $\mathbf{x}_i$.

The ATE is estimated by replacing (14f) with

$$1/N \sum_{i=1}^{N} \varpi_i \left\{ \widehat{E}(t_i|\mathbf{x}_i, \tau = 1, \widehat{\boldsymbol{\beta}}_{\text{ipwrar,ate},1}) - \widehat{\text{POM}}_{\text{ipwrar},0} - \widehat{\text{ATE}}_{\text{ipwrar}} \right\} = 0$$

For the ATET, the contribution of the ith observation to the weighted log likelihood that is maximized in step IPWRAR3 is

$$L_{\text{ipwrar}}(t_i, \mathbf{x}_i, \tau, \widehat{\boldsymbol{\beta}}_{\text{ipwrar,ate},\tau}) = \varpi_i \omega_{i,\text{cot},\tau} (\tau_i == \tau) \ln\{ f(t_i|\mathbf{x}_i, \tau, \widehat{\boldsymbol{\beta}}_{\text{ipwrar,atet},\tau}) \}$$

where $\omega_{i,\text{cot},1}$ is given in (10a), $\omega_{i,\text{cot},0}$ is given in (10b), and $f(t_i|\mathbf{x}_i, \tau, \widehat{\boldsymbol{\beta}}_{\text{ipwrar,atet},\tau})$ is the density corresponding to distribution $F(t_i|\mathbf{x}_i, \tau, \widehat{\boldsymbol{\beta}}_{\text{ipwrar,atet},\tau})$.

The WAC-IPWRA estimators for the conditional-on-treatment POMs simultaneously solve estimating equations (15a) through (15f) for $\widehat{\boldsymbol{\beta}}_{\text{ipwrar,atet},0}$, $\widehat{\boldsymbol{\beta}}_{\text{ipwrar,atet},0}$, $\widehat{\boldsymbol{\gamma}}$, $\widehat{\boldsymbol{\alpha}}$, $\widehat{\text{POM}}_{\text{ipwrar,cot},0}$, and $\widehat{\text{POM}}_{\text{ipwrar,cot},1}$.

$$1/N \sum_{i=1}^{N} \mathbf{s}_{\text{ipwrar}}(t_i, \mathbf{w}_i, \widehat{\boldsymbol{\gamma}}, F_c) = \mathbf{0} \tag{15a}$$

$$1/N \sum_{i=1}^{N} \mathbf{s}_{\text{ipwrar}}(\tau_i, \mathbf{z}_i, \widehat{\boldsymbol{\alpha}}, p) = \mathbf{0} \tag{15b}$$

$$1/N \sum_{i=1}^{N} \mathbf{s}_{\text{ipwrar}}(t_i, \mathbf{x}_i, 0, \widehat{\boldsymbol{\beta}}_{\text{ipwrar,atet},0}, F) = \mathbf{0} \tag{15c}$$

$$1/N \sum_{i=1}^{N} \mathbf{s}_{\text{ipwrar}}(t_i, \mathbf{x}_i, 1, \widehat{\boldsymbol{\beta}}_{\text{ipwrar,atet},1}, F) = \mathbf{0} \tag{15d}$$

$$1/N \sum_{i=1}^{N} \varpi_i \left\{ \widehat{E}(t_i|\mathbf{x}_i, \tau = 0, \widehat{\boldsymbol{\beta}}_{\text{ipwrar,atet},0}) - \widehat{\text{POM}}_{\text{ipwrar,cot},0} \right\} = 0 \tag{15e}$$

$$1/N \sum_{i=1}^{N} \varpi_i \left\{ \widehat{E}(t_i|\mathbf{x}_i, \tau = 1, \widehat{\boldsymbol{\beta}}_{\text{ipwrar,atet},1}) - \widehat{\text{POM}}_{\text{ipwrar,cot},1} \right\} = 0 \tag{15f}$$

where

$\mathbf{s}_{\text{ipwrar}}(t_i, \mathbf{w}_i, \widehat{\boldsymbol{\gamma}}, F_c) = \frac{\partial L_{c,\text{ipwrar}}(t_i, \mathbf{w}_i, \widehat{\boldsymbol{\gamma}})}{\partial \widehat{\boldsymbol{\gamma}}}$ is the vector of score equations from the ML estimator for $\widehat{\boldsymbol{\gamma}}$ based on survival-time model F_c,

$\mathbf{s}_{\text{ipwrar}}(\tau_i, \mathbf{z}_i, \widehat{\boldsymbol{\alpha}}, p) = \frac{\partial L_{p,\text{ipwrar}}(\tau_i, \mathbf{z}_i, \widehat{\boldsymbol{\alpha}})}{\partial \widehat{\boldsymbol{\alpha}}}$ is the vector of score equations from the ML estimator for $\widehat{\boldsymbol{\alpha}}$ based on probability model p,

$\mathbf{s}_{\text{ipwrar}}(t_i, \mathbf{x}_i, 0, \widehat{\boldsymbol{\beta}}_{\text{ipwrar,atet,0}}, F) = \frac{\partial L_{\text{ipwrar}}(t_i, \mathbf{x}_i, 0, \widehat{\boldsymbol{\beta}}_{\text{ipwrar,atet,0}})}{\partial \widehat{\boldsymbol{\beta}}_{\text{ipwrar,atet,0}}}$ is the vector of score equations from the WML estimator for $\widehat{\boldsymbol{\beta}}_{\text{ipwrar,atet,0}}$ based on survival-time model F,

$\mathbf{s}_{\text{ipwrar}}(t_i, \mathbf{x}_i, 1, \widehat{\boldsymbol{\beta}}_{\text{ipwrar,atet,1}}, F) = \frac{\partial L_{\text{ipwrar}}(t_i, \mathbf{x}_i, 1, \widehat{\boldsymbol{\beta}}_{\text{ipwrar,atet,1}})}{\partial \widehat{\boldsymbol{\beta}}_{\text{ipwrar,atet,1}}}$ is the vector of score equations from the WML estimator for $\widehat{\boldsymbol{\beta}}_{\text{ipwrar,atet,1}}$ based on survival-time model F,

$\widehat{E}(t_i|\mathbf{x}_i, \tau = 0, \widehat{\boldsymbol{\beta}}_{\text{ipwrar,atet,0}})$ is the predicted mean survival time assuming treatment level 0 for observation i conditional on $\mathbf{x}_i$, and

$\widehat{E}(t_i|\mathbf{x}_i, \tau = 1, \widehat{\boldsymbol{\beta}}_{\text{ipwrar,atet,1}})$ is the predicted mean survival time assuming treatment level 1 for observation i conditional on $\mathbf{x}_i$.

The ATET is estimated by replacing (15f) with

$$1/N \sum_{i=1}^{N} \varpi_i \left\{ \widehat{E}(t_i|\mathbf{x}_i, \tau = 1, \widehat{\boldsymbol{\beta}}_{\text{ipwrar,atet,1}}) - \widehat{\text{POM}}_{\text{ipwrar,cot,0}} - \widehat{\text{ATET}}_{\text{ipwrar}} \right\} = 0$$

Likelihood-adjusted-censoring IPWRA

When a censoring model is not specified, `stteffects ipwra` uses the formulas for the LAC-IPWRA estimator that add a term to the log-likelihood function. For notational conciseness and to reinforce its use of an additional term in the log likelihood, we denote the LAC-IPWRA estimator by IPWRAL in lists and formulas.

The methods and formulas for the LAC-IPWRA estimator differ in three ways from those for the WAC-IPWRA estimator.

1. No censoring model is specified, so LAC-IPWRA does not perform a version of step IPWRAR1 and it does not use the moment equations (14a).
2. The weights only depend on the treatment level and treatment assignment probabilities, not on the censoring.
3. The WML estimator for $\boldsymbol{\beta}_\tau$ includes a term for censored observations and censored observations are used. Recall that for the WAC-IPWRA estimator, the weights used in the WML estimator for $\boldsymbol{\beta}_\tau$ account for the censoring, and the censored observations are not used in the WML estimator.

The LAC-IPWRA estimators have the following logic.

IPWRAL1. Estimate by ML the parameters $\boldsymbol{\alpha}$ of a parametric model for the probability of treatment model $p(\mathbf{z}_i, \boldsymbol{\alpha})$.

IPWRAL2. For each treatment level $\tau \in \{0, 1\}$, estimate by WML the parameters $\boldsymbol{\beta}_\tau$ of a parametric model for the survival-time outcome t in which $F(t|\mathbf{x}, \tau, \boldsymbol{\beta}_\tau)$ is the distribution of t conditional on covariates $\mathbf{x}$ and treatment level τ. The weights depend only on the treatment level and the treatment-assignment probabilities. For the ATE, the weights are those in (12a) and (12b). For the ATET, the weights are those in (13a) and (13b). We denote the estimates of $\boldsymbol{\beta}_\tau$ by $\widehat{\boldsymbol{\beta}}_{\text{ipwral},\tau}$.

IPWRAL3. Use the estimated $\widehat{\boldsymbol{\beta}}_{\text{ipwral},\tau}$ and the functional form implied by $F(t|\mathbf{x},\tau,\boldsymbol{\beta}_\tau)$ to estimate the mean survival time, conditional on $\mathbf{x}$ and treatment level τ, for each sample observation, denoted by $\widehat{E}(t_i|\mathbf{x}_i,\tau,\widehat{\boldsymbol{\beta}}_{\text{ipwral},\tau})$. Conditional independence of the treatment and the survival-time potential outcomes ensures that $E(t|\mathbf{x},\tau,\boldsymbol{\beta}_\tau) = E(t_\tau|\mathbf{x},\boldsymbol{\beta}_\tau)$, where t_τ is the potential survival-time outcome corresponding to treatment level τ. Under correct model specification, sample averages of $\widehat{E}(t_i|\mathbf{x}_i,\tau,\widehat{\boldsymbol{\beta}}_{\text{ipwral},\tau})$ consistently estimate the POM for treatment level τ, denoted by POM_τ.

The contribution of the ith observation to the log likelihood that is maximized in step IPWRAL1 is

$$L_{p,\text{ipwral}}(\tau_i,\mathbf{z}_i,\widehat{\boldsymbol{\alpha}}) = \varpi_i\left[(\tau_i == 1)\ln\{p(\mathbf{z}_i,\widehat{\boldsymbol{\alpha}})\} + \{1-(\tau_i == 1)\}\ln\{1-p(\mathbf{z}_i,\widehat{\boldsymbol{\alpha}})\}\right]$$

where $p(\mathbf{z}_i,\widehat{\boldsymbol{\alpha}})$ is the model for the probability that i gets treatment level 1.

The weights and the parameters in step IPWRAL2 used to estimate the ATE differ from those used to estimate the ATET. For the ATE, the contribution of the ith observation to the log likelihood that is maximized in step IPWRAL2 is

$$L_{\text{ipwral}}(t_i,\mathbf{x}_i,\tau,\widehat{\boldsymbol{\beta}}_{\text{ipwral,ate},\tau}) = (\tau_i == \tau)\varpi_i\omega_{i,\tau}\Big\{(1-c_i)\ln\{f(t_i|\mathbf{x}_i,\tau,\widehat{\boldsymbol{\beta}}_{\text{ipwrar,ate},\tau})\} \\ c_i\ln\{1-F(t_i|\mathbf{x}_i,\tau,\widehat{\boldsymbol{\beta}}_{\text{ipwrar,ate},\tau})\}\Big\}$$

where $\omega_{i,1}$ is given in (12a), $\omega_{i,0}$ is given in (12b), and $f(t_i|\mathbf{x}_i,\tau,\widehat{\boldsymbol{\beta}}_{\text{ipwral,ate},\tau})$ is the density corresponding to distribution $F(t_i|\mathbf{x}_i,\tau,\widehat{\boldsymbol{\beta}}_{\text{ipwral,ate},\tau})$. Unlike the WRA estimator, the censored observations are used, and there is a term in the likelihood function to account for censoring.

The LAC-IPWRA estimators for the POMs simultaneously solve estimating equations (16a) through (16e) for $\widehat{\boldsymbol{\alpha}}$, $\widehat{\boldsymbol{\beta}}_{\text{ipwral,ate},0}$, $\widehat{\boldsymbol{\beta}}_{\text{ipwral,ate},0}$, $\widehat{\text{POM}}_{\text{ipwral},0}$, and $\widehat{\text{POM}}_{\text{ipwral},1}$.

$$1/N\sum_{i=1}^{N}\mathbf{s}_{\text{ipwral}}(\tau_i,\mathbf{z}_i,\widehat{\boldsymbol{\alpha}},p) = \mathbf{0} \tag{16a}$$

$$1/N\sum_{i=1}^{N}\mathbf{s}_{\text{ipwral}}(t_i,\mathbf{x}_i,0,\widehat{\boldsymbol{\beta}}_{\text{ipwral,ate},0},F) = \mathbf{0} \tag{16b}$$

$$1/N\sum_{i=1}^{N}\mathbf{s}_{\text{ipwral}}(t_i,\mathbf{x}_i,1,\widehat{\boldsymbol{\beta}}_{\text{ipwral,ate},1},F) = \mathbf{0} \tag{16c}$$

$$1/N\sum_{i=1}^{N}\varpi_i\left\{\widehat{E}(t_i|\mathbf{x}_i,\tau=0,\widehat{\boldsymbol{\beta}}_{\text{ipwral,ate},0}) - \widehat{\text{POM}}_{\text{ipwral},0}\right\} = 0 \tag{16d}$$

$$1/N\sum_{i=1}^{N}\varpi_i\left\{\widehat{E}(t_i|\mathbf{x}_i,\tau=1,\widehat{\boldsymbol{\beta}}_{\text{ipwral,ate},1}) - \widehat{\text{POM}}_{\text{ipwral},1}\right\} = 0 \tag{16e}$$

where

$\mathbf{s}_{\text{ipwral}}(\tau_i,\mathbf{z}_i,\widehat{\boldsymbol{\alpha}},p) = \frac{\partial L_{p,\text{ipwral}}(\tau_i,\mathbf{z}_i,\widehat{\boldsymbol{\alpha}})}{\partial\widehat{\boldsymbol{\alpha}}}$ is the vector of score equations from the ML estimator for $\widehat{\boldsymbol{\alpha}}$ based on probability model p,

$\mathbf{s}_{\text{ipwral}}(t_i,\mathbf{x}_i,0,\widehat{\boldsymbol{\beta}}_{\text{ipwral,ate,0}},F) = \frac{\partial L_{\text{ipwral}}(t_i,\mathbf{x}_i,0,\widehat{\boldsymbol{\beta}}_{\text{ipwral,ate,0}})}{\partial\widehat{\boldsymbol{\beta}}_{\text{ipwral,ate,0}}}$ is the vector of score equations from the WML estimator for $\widehat{\boldsymbol{\beta}}_{\text{ipwral,ate,0}}$ based on survival-time model F,

$\mathbf{s}_{\text{ipwral}}(t_i,\mathbf{x}_i,1,\widehat{\boldsymbol{\beta}}_{\text{ipwral,ate,1}},F) = \frac{\partial L_{\text{ipwral}}(t_i,\mathbf{x}_i,1,\widehat{\boldsymbol{\beta}}_{\text{ipwral,ate,1}})}{\partial\widehat{\boldsymbol{\beta}}_{\text{ipwral,ate,1}}}$ is the vector of score equations from the WML estimator for $\widehat{\boldsymbol{\beta}}_{\text{ipwral,ate,1}}$ based on survival-time model F,

$\widehat{E}(t_i|\mathbf{x}_i,\tau=0,\widehat{\boldsymbol{\beta}}_{\text{ipwral,ate,0}})$ is the predicted mean survival time assuming treatment level 0 for observation i conditional on $\mathbf{x}_i$, and

$\widehat{E}(t_i|\mathbf{x}_i,\tau=1,\widehat{\boldsymbol{\beta}}_{\text{ipwral,ate,1}})$ is the predicted mean survival time assuming treatment level 1 for observation i conditional on $\mathbf{x}_i$.

The ATE is estimated by replacing (16e) with

$$1/N\sum_{i=1}^{N}\varpi_i\left\{\widehat{E}(t_i|\mathbf{x}_i,\tau=1,\widehat{\boldsymbol{\beta}}_{\text{ipwral,ate,1}}) - \widehat{\text{POM}}_{\text{ipwral,0}} - \widehat{\text{ATE}}_{\text{ipwral}}\right\} = 0$$

For the ATET, the contribution of the ith observation to the WML function that is maximized in step IPWRAL2 is

$$L_{\text{ipwral}}(t_i,\mathbf{x}_i,\tau,\widehat{\boldsymbol{\beta}}_{\text{ipwral,atet},\tau}) = (\tau_i == \tau)\varpi_i\omega_{i,\text{cot},\tau}\left\{(1-c_i)\ln\{f(t_i|\mathbf{x}_i,\tau,\widehat{\boldsymbol{\beta}}_{\text{ipwrar,atet},\tau})\}\right.$$
$$\left. c_i\ln\{1-F(t_i|\mathbf{x}_i,\tau,\widehat{\boldsymbol{\beta}}_{\text{ipwrar,atet},\tau})\}\right\}$$

where $\omega_{i,\text{cot},1}$ is given in (13a), $\omega_{i,\text{cot},0}$ is given in (13b), and $f(t_i|\mathbf{x}_i,\tau,\widehat{\boldsymbol{\beta}}_{\text{ipwral,atet},\tau})$ is the density corresponding to distribution $F(t_i|\mathbf{x}_i,\tau,\widehat{\boldsymbol{\beta}}_{\text{ipwral,atet},\tau})$. Again unlike the WRA, the censored observations are used, and there is a term in the likelihood function to account for censoring.

The LAC-IPWRA estimators for the conditional-on-treatment POMs simultaneously solve estimating equations (17a) through (17e) for $\widehat{\boldsymbol{\alpha}}$, $\widehat{\boldsymbol{\beta}}_{\text{ipwral,atet,0}}$, $\widehat{\boldsymbol{\beta}}_{\text{ipwral,atet,0}}$, $\widehat{\text{POM}}_{\text{ipwral,cot,0}}$, and $\widehat{\text{POM}}_{\text{ipwral,cot,1}}$.

$$1/N\sum_{i=1}^{N}\mathbf{s}_{\text{ipwral}}(\tau_i,\mathbf{z}_i,\widehat{\boldsymbol{\alpha}},p) = \mathbf{0} \tag{17a}$$

$$1/N\sum_{i=1}^{N}\mathbf{s}_{\text{ipwral}}(t_i,\mathbf{x}_i,0,\widehat{\boldsymbol{\beta}}_{\text{ipwral,atet,0}},F) = \mathbf{0} \tag{17b}$$

$$1/N\sum_{i=1}^{N}\mathbf{s}_{\text{ipwral}}(t_i,\mathbf{x}_i,1,\widehat{\boldsymbol{\beta}}_{\text{ipwral,atet,1}},F) = \mathbf{0} \tag{17c}$$

$$1/N\sum_{i=1}^{N}\varpi_i\left\{\widehat{E}(t_i|\mathbf{x}_i,\tau=0,\widehat{\boldsymbol{\beta}}_{\text{ipwral,atet,0}}) - \widehat{\text{POM}}_{\text{ipwral,cot,0}}\right\} = 0 \tag{17d}$$

$$1/N\sum_{i=1}^{N}\varpi_i\left\{\widehat{E}(t_i|\mathbf{x}_i,\tau=1,\widehat{\boldsymbol{\beta}}_{\text{ipwral,atet,1}}) - \widehat{\text{POM}}_{\text{ipwral,cot,1}}\right\} = 0 \tag{17e}$$

where

$\mathbf{s}_{\text{ipwral}}(\tau_i, \mathbf{z}_i, \widehat{\boldsymbol{\alpha}}, p) = \frac{\partial L_{p,\text{ipwral}}(\tau_i, \mathbf{z}_i, \widehat{\boldsymbol{\alpha}})}{\partial \widehat{\boldsymbol{\alpha}}}$ is the vector of score equations from the ML estimator for $\widehat{\boldsymbol{\alpha}}$ based on probability model p,

$\mathbf{s}_{\text{ipwral}}(t_i, \mathbf{x}_i, 0, \widehat{\boldsymbol{\beta}}_{\text{ipwral,atet,0}}, F) = \frac{\partial L_{\text{ipwral}}(t_i, \mathbf{x}_i, 0, \widehat{\boldsymbol{\beta}}_{\text{ipwral,atet,0}})}{\partial \widehat{\boldsymbol{\beta}}_{\text{ipwral,atet,0}}}$ is the vector of score equations from the WML estimator for $\widehat{\boldsymbol{\beta}}_{\text{ipwral,atet,0}}$ based on survival-time model F,

$\mathbf{s}_{\text{ipwral}}(t_i, \mathbf{x}_i, 1, \widehat{\boldsymbol{\beta}}_{\text{ipwral,atet,1}}, F) = \frac{\partial L_{\text{ipwral}}(t_i, \mathbf{x}_i, 1, \widehat{\boldsymbol{\beta}}_{\text{ipwral,atet,1}})}{\partial \widehat{\boldsymbol{\beta}}_{\text{ipwral,atet,1}}}$ is the vector of score equations from the WML estimator for $\widehat{\boldsymbol{\beta}}_{\text{ipwral,atet,1}}$ based on survival-time model F,

$\widehat{E}(t_i|\mathbf{x}_i, \tau = 0, \widehat{\boldsymbol{\beta}}_{\text{ipwral,atet,0}})$ is the predicted mean survival time assuming treatment level 0 for observation i conditional on $\mathbf{x}_i$, and

$\widehat{E}(t_i|\mathbf{x}_i, \tau = 1, \widehat{\boldsymbol{\beta}}_{\text{ipwral,atet,1}})$ is the predicted mean survival time assuming treatment level 1 for observation i conditional on $\mathbf{x}_i$.

The ATET is estimated by replacing (17e) with

$$1/N \sum_{i=1}^{N} \varpi_i \left\{ \widehat{E}(t_i|\mathbf{x}_i, \tau = 1, \widehat{\boldsymbol{\beta}}_{\text{ipwral,atet,1}}) - \widehat{\text{POM}}_{\text{ipwral,cot,0}} - \widehat{\text{ATET}}_{\text{ipwral}} \right\} = 0$$

Functional-form details

In this section, we specify the functional forms for the conditional distribution function used in the survival-time outcome model F, the conditional distribution function used in the survival-time censoring model F_c, and the conditional distribution used to model the treatment probabilities p.

You may choose among the same set of conditional distribution functions for either F or F_c: `exponential`, `weibull`, `lnormal`, or `gamma`.

Name	Cumulative	Density	Mean
exponential	$1 - \exp(-\lambda_i t_i)$	$\lambda_i \exp(-\lambda_i t_i)$	$1/\lambda_i$
Weibull	$1 - \exp\{-(\lambda_i t_i)^{s_i}\}$	$s_i t_i^{s_i - 1} \lambda_i^{s_i} \exp\{-(\lambda_i t_i)^{s_i}\}$	$(1/\lambda_i)\Gamma\{(s_i + 1)/s_i\}$
log normal	$\Phi\{(\ln(t_i) - \lambda_i)/s_i\}$	$(1/(s_i t_i))\phi\{(\ln(t_i) - \lambda_i)/s_i\}$	$\exp(\lambda_i + s_i^2/2)$
gamma	$\texttt{gammap}\{s_i, (s_i t_i/\lambda_i)\}$	$(s_i^{s_i} t_i^{s_i - 1})/\{\lambda_i^{s_i}\Gamma(s_i)\}\exp(-s_i t_i/\lambda_i)$	λ_i

where the following table specifies how λ_i and s_i are parameterized in terms of the covariates $\mathbf{x}_i$ and the ancillary covariates $\widetilde{\mathbf{x}}_i$, respectively.

Name	λ_i	s_i
exponential	$\exp(-\mathbf{x}_i\boldsymbol{\beta})$	
Weibull	$\exp(-\mathbf{x}_i\boldsymbol{\beta})$	$\exp(\widetilde{\mathbf{x}}_i\widetilde{\boldsymbol{\beta}})$
log normal	$\mathbf{x}_i\boldsymbol{\beta}$	$\exp(\widetilde{\mathbf{x}}_i\widetilde{\boldsymbol{\beta}})$
gamma	$\exp(\mathbf{x}_i\boldsymbol{\beta})$	$\exp(-2\widetilde{\mathbf{x}}_i\widetilde{\boldsymbol{\beta}})$

For the treatment-assignment models, the `probit` model uses the standard normal distribution, the `logit` uses the standard logistic distribution, the `hetprobit` model uses

$$\Phi\{\mathbf{z}_1\boldsymbol{\alpha}_1/\exp(\mathbf{z}_2\boldsymbol{\alpha}_2)\}$$

and the multinomial logit uses

$$p(\mathbf{z},t) = \exp(\mathbf{z}\boldsymbol{\alpha}_t)/\{1+\sum_{k=1}^{q}\exp(\mathbf{z}\boldsymbol{\alpha}_k)\}$$

where the notation is defined below.

In the `hetprobit` model, $\mathbf{z}_1$ are the covariates specified in the treatment-assignment specification, $\mathbf{z}_2$ are the covariates specified in the `hetprobit()` option, and $\boldsymbol{\alpha}_1$ and $\boldsymbol{\alpha}_2$ are the corresponding coefficients.

In the multinomial logit model, $\mathbf{z}$ are the covariates specified in the treatment-assignment specification and α_k are the coefficients; see [R] **mlogit** for further details.

References

Angrist, J. D., and J.-S. Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press.

Bai, X., A. A. Tsiatis, and S. M. O'Brien. 2013. Doubly robust estimators of treatment-specific survival distributions in observational studies with stratified sampling. *Biometrics* 69: 830–839. https://doi.org/10.1111/biom.12076.

Busso, M., J. DiNardo, and J. McCrary. 2014. New evidence on the finite sample properties of propensity score reweighting and matching estimators. *Review of Economics and Statistics* 96: 885–897. https://doi.org/10.1162/REST_a_00431.

Cameron, A. C., and P. K. Trivedi. 2005. *Microeconometrics: Methods and Applications*. New York: Cambridge University Press.

Cattaneo, M. D. 2010. Efficient semiparametric estimation of multi-valued treatment effects under ignorability. *Journal of Econometrics* 155: 138–154. https://doi.org/10.1016/j.jeconom.2009.09.023.

Cattaneo, M. D., D. M. Drukker, and A. D. Holland. 2013. Estimation of multivalued treatment effects under conditional independence. *Stata Journal* 13: 407–450.

Drukker, D. M. 2014. Using gmm to solve two-step estimation problems. *The Stata Blog: Not Elsewhere Classified*. http://blog.stata.com/2014/12/08/using-gmm-to-solve-two-step-estimation-problems/.

Guo, S., and M. W. Fraser. 2015. *Propensity Score Analysis: Statistical Methods and Applications*. 2nd ed. Thousand Oaks, CA: Sage.

Hirano, K., G. W. Imbens, and G. Ridder. 2003. Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica* 71: 1161–1189. https://doi.org/10.1111/1468-0262.00442.

Imbens, G. W. 2000. The role of the propensity score in estimating dose–response functions. *Biometrika* 87: 706–710. https://doi.org/10.1093/biomet/87.3.706.

——. 2004. Nonparametric estimation of average treatment effects under exogeneity: A review. *Review of Economics and Statistics* 86: 4–29. https://doi.org/10.1162/003465304323023651.

Imbens, G. W., and J. M. Wooldridge. 2009. Recent developments in the econometrics of program evaluation. *Journal of Economic Literature* 47: 5–86. https://doi.org/10.1257/jel.47.1.5.

Lin, D. Y. 2000. Linear regression analysis of censored medical costs. *Biostatistics* 1: 35–47. https://doi.org/10.1093/biostatistics/1.1.35.

Newey, W. K. 1984. A method of moments interpretation of sequential estimators. *Economics Letters* 14: 201–206. https://doi.org/10.1016/0165-1765(84)90083-1.

Newey, W. K., and D. L. McFadden. 1994. Large sample estimation and hypothesis testing. In Vol. 4 of *Handbook of Econometrics*, ed. R. F. Engle and D. L. McFadden, 2111–2245. Amsterdam: Elsevier. https://doi.org/10.1016/S1573-4412(05)80005-4.

Robins, J. M., and A. Rotnitzky. 1995. Semiparametric efficiency in multivariate regression models with missing data. *Journal of the American Statistical Association* 90: 122–129. https://doi.org/10.2307/2291135.

——. 2006. Inverse probability weighting in survival analysis. In *Survival and Event History Analysis*, ed. N. Keiding and P. K. Andersen, 266–271. Chichester, UK: Wiley.

Robins, J. M., A. Rotnitzky, and L. P. Zhao. 1995. Analysis of semiparametric regression models for repeated outcomes in the presence of missing data. *Journal of the American Statistical Association* 90: 106–121. https://doi.org/10.1080/01621459.1995.10476493.

Rosenbaum, P. R., and D. B. Rubin. 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika* 70: 41–55. https://doi.org/10.2307/2335942.

Rubin, D. B. 1974. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology* 66: 688–701. https://doi.org/10.1037/h0037350.

Tsiatis, A. A. 2006. *Semiparametric Theory and Missing Data*. New York: Springer.

Vittinghoff, E., D. V. Glidden, S. C. Shiboski, and C. E. McCulloch. 2012. *Regression Methods in Biostatistics: Linear, Logistic, Survival, and Repeated Measures Models*. 2nd ed. New York: Springer.

Wooldridge, J. M. 2002. Inverse probability weighted M-estimators for sample selection, attrition, and stratification. *Portuguese Economic Journal* 1: 117–139. https://doi.org/10.1007/s10258-002-0008-x.

——. 2007. Inverse probability weighted estimation for general missing data problems. *Journal of Econometrics* 141: 1281–1301. https://doi.org/10.1016/j.jeconom.2007.02.002.

——. 2010. *Econometric Analysis of Cross Section and Panel Data*. 2nd ed. Cambridge, MA: MIT Press.

Also see

[CAUSAL] **stteffects postestimation** — Postestimation tools for stteffects

[CAUSAL] **stteffects intro** — Introduction to treatment effects for observational survival-time data

[ST] **streg** — Parametric survival models

[ST] **stset** — Declare data to be survival-time data

[U] 20 Estimation and postestimation commands